# Predicting of Surface Ozone Using Artificial Neural Networks and Support Vector Machines

Mouhammd Alkasassbeh

*Mutah University, Jordan*
*(malkasasbeh@gmail.com)*

***Abstract***

*Due to increase in industrial and anthropogenic activities, air pollution has been a serious environmental problem all over the world. It was found that harmful emission into the air is a symbol for environmental force that affects seriously man's health, natural life and agriculture; thus leading to major loss of the nation's economy. In this paper, the prediction of the surface ozone layer problem is explored. A comparison between two types of Artificial Neural Networks (ANN) (i.e. back propagation and Radial Basis Functions (RBF) networks) and the Support Vector Machines (SVM) techniques for short prediction of surface ozone is conclusively demonstrated. Three models which predict the expected values of the surface ozone based on three variables (i.e. Nitrogen-di-oxide, temperature and Relative Humidity) will be presented.*

***Keywords:*** *Air pollution; Surface ozone; back propagation neural network; Radial Basis Function (RBF) neural network; Support Vector Machines*

## 1. Introduction

Ozone Layer is part of the atmosphere around the planet which has extensive ozone gas (O3). It is largely centered in the lower part of the stratosphere of the Earth's atmosphere [1]. Moreover, the ozone layer is surrounding the atmosphere fully briefing at an altitude between 20 and 30 kilometers from the Earth's surface and with thickness ranges of 2-8 kilometers. Correspondingly, the ozone layer role natural filter and protective shield that surrounds the earth to protect it from the harmful Ultra Violet (UV) serious damages to human health and animal alike [1–3], and also reduces plant growth and crop production, and affects the systems aquatic environment [4]. Human activities and new technology in the synthesis of chemicals have led to the destruction of the ozone layer. One of the most important chemicals that depletes the ozone layer is the chlorofluorocarbon (cfc) which is frequently used in refrigeration and air-conditioning equipment [5]. The methyl bromide is also used as a pesticide in stored agricultural crops and agricultural soil sterilization [6]. Likewise, some of the solvents used in cleaning metal mechanical parts and electronic circuits such as carbon tetrachloride material.

The environmental damage caused by erosion of the ozone layer is substantially in the climatic changes incident to the planet, including the abrupt changes in weather and climate [7] , desertification and forest fires and the rise in sea level to the shores of many in the world and disrupt the ecological balance and health damage is in some skin cancers and immune suppression natural for humans and some eye diseases such as the opacity of the eye if exposure of living organisms, such as humans and animals to large amounts of the harmful UV.

The destruction of the ozone layer is the biggest examples of the danger that all humans facing with all their different circumstances environmental. The international community considered and started to enforce it by the purely legal defined problems. For prediction, developments of mathematical models are clearly beneficial in many areas like prediction of prices in the stock market [8, 9], earthquake prediction [10], prediction of fermentation [11], stream flood simulation and prediction [12] and river flow prediction [13, 14]. Different approaches have been developed for solving the ozone predication problem. For example, Auto-Regressive Moving Average (ARMA) [15] was used in the forecasting of ambient air pollutants. Authors show that ARMA model can be utilized for the air quality forewarning purposes. ANN was used to develop a prediction model for the SO2 [16]. Recently, prediction of SO2 using FF-ANN was presented in [17]. Author showed that Multi-Layer Perceptron (MLP) can provide better results than linear regression technique. The model can help in strategic planning and management of local air quality. ANN additionally was used for short term prediction of surface ozone in [18]. Authors claim that the developed model can perform satisfactory well. However, they did not give details on the training, testing or even the validation data set, not either the structure of the developed ANN and its parameters. In this paper, different well-known machine learning methods have been studied for predation of the surface ozone; the main objective of this prediction is to give an instant alert before anything uncontrolled will occur. In this work, we provide three models for the prediction of the surface ozone layer. Back propagation ANN, RBF-ANN and SVM were used to build a suitable model which can be used to solve the prediction problem. In section 2, we present the collected data set from [18].

A brief introduction to the two types of network under study in this paper is presented in section III. In section IV, we give an introduction to Support Vector Machines. The way we measure the performance of the developed models is presented in section V. Detailed results are presented in section VI. Finally, we provide our conclusions.

## 2. Area of Study and Data Description

The study area is Chenbagaramanputhur (8o15’1”N, 77o29’19”E)(see Figure 1), it is a countryside in Kanyakumari district; it is about 12 kms from Nagercoil town. The operating temperature range is from $5^{o}$C to $50^{o}$C, relative humidity limits are 5% and 95%. Similar kind of NO2 sensor has been used for nitrogen dioxide measurement, the gas sensitive semiconductor (GSS) type sensor is described in (www.aeroqual.com). The sampling was carried out for three months from May 2009 to July 2009. For ozone, seven readings were taken per day starting from 530h to 2330h with three hour interval. For NO2, only two readings were taken one at daytime and the other at night time. Furthermore, the data set, used in this work, carried out for three months from May 2009 to July 2009. For the ozone, seven readings were taken per with three hour interval. For the NO2, only two readings were taken each day, one at daytime and the other at nighttime. The model inputs and Output are given in Table I. Figure 2 shows the collected measurements to develop the model as given in [18].

**Table I.**

Inputs and output for the surface ozone model

| | | |
|---|---|---|
| Inputs | Nitrogen dioxide concentration | $x_1$ |
| | Mean temperature | $x_2$ |
| | Prevailing % Relative Humidity | $x_3$ |
| Output | Mean surface ozone concentration | $y$ |


**Figure 1. Location of the Area of Study**

## 3. Artificial Neural Network

Artificial neural networks (ANNs) were introduced by Mcculloch and Pitts in 1943 [19]. They represent parallel working systems, which consist of a number of processing units interconnected in a network, these units are called neurons. ANNs consist of a number of elements called neurons. These neurons are grouped together to form a layer. Each neuron has a number of inputs and a single output. Each input has an assigned factor or parameter called the weight. The way how a neuron is working is as follows: input signal to each neuron is multiplication by the corresponding weight then the result from the multiplication is summed and passes through a transfer function, most likely to be a sigmoid function. If the result from the summation is over a certain threshold, the neuron output will be activated else the output is not.

There are two main features make neural networks a very useful tool in solving prediction and modeling problems. Those features are; their ability to learn and generalize. Learning process depend on providing a set of training data, which used to adjust the network weights, using a described learning algorithm. After training process have been successfully achieved, the network will be able to recognize a certain output when a newly data is presented to its input layer. This is what we mean by generalization. Finding the network weights, such that the difference between the network output and the desired output; is the main job of the ANN learning algorithm. ANNs have been used both to estimate parameters of a formal model and to learn to emulate the process model itself to predict future outcomes. ANNs have been successfully used to solve a variety of prediction and forecasting problems. It has proven to be efficient in a number of applications, including predicting sales [20], forecasting prices [21], predicting shift failures [22] and predicting stock returns [23]. It was also used for the prediction of river flows [13].

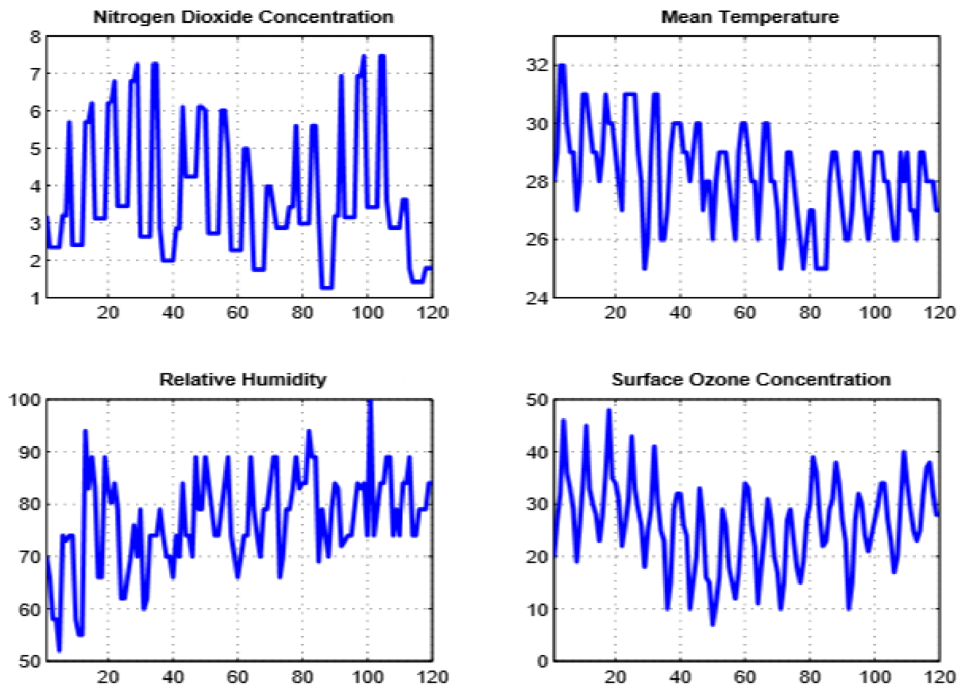


**Figure 2. Training and Testing Data Set [18]**

### 3.1. FF-ANN

In this paper, the first model that we adopted is the feed forward (FF) fully connected networks. The structure of this type of network is presented in Figure 3. Moreover, it allows the data flow to travel from one direction only from input to output. There is no feedback; it tends to be straight forward networks that associate inputs with outputs.

ANNs can be distinguished by a number of performance characteristics, which can be summarized into three points [24, 25]:

- **Neural Network architecture**: The pattern of connections between the neurons in different layers. An example of simple ANN architecture is shown in Figure 3.
- **Training Algorithm**: it the way to determine the weights of the connections in order to minimize the difference between the actual and predicted value as much as possible.
- **Activation Function**: Activation function is applied by each neuron to its net input in two stages, (sum of weight input signals, Equation 2) to determine its output signal. This function is usually nonlinear. Sigmoid function (S-shaped curve) is one of the most common used activation functions. The sigmoid function can be given by the Equation 2.

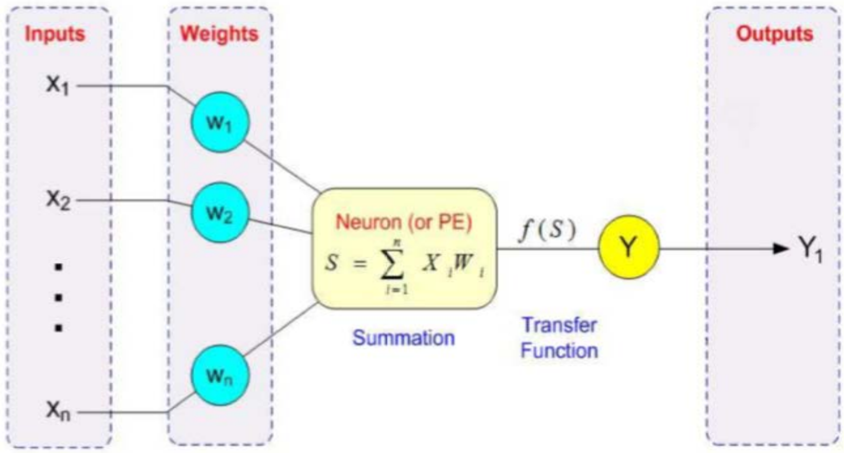


**Figure 3. FF-ANN Structure**

$$S = \sum_{i=1}^{n} x_i w_i \qquad (1)$$

$$f(S) = \frac{1}{1+e^{-S}} \qquad (2)$$

### 3.2. RBF-ANN

RBF Network is composed of two-layer feed forward neural network. Processing units of the network are radial functions given by Equation 3 for $i = 1,..,L$ where $X$ is the input feature vector, $L$ is the number of hidden layers, $\mu_i$ and $\sum i$ are the mean and covariance matrix of the *ith* Gaussian function.

$$\phi_i(X) = exp[-(X-\mu_i)^T \sum_i^{-1} (X-\mu_i)] \qquad (3)$$

RBF networks are different on emphasizing the training part as it retraining as much as possible therefore the training is fairly fast. This emphasizing gives the network very simple mathematics, it is basically linear algebra. RBF network typically has three layers: an input layer, a hidden layer with a nonlinear RBF activation function (usually Gaussian function) and an output layer which implements a linear activation functions as shown in Figure 4. RBF is used mainly for regression and for performing complex pattern classification tasks. In this work, RBF shows good results and performance. It supersedes the FF-ANN as given in Table V.

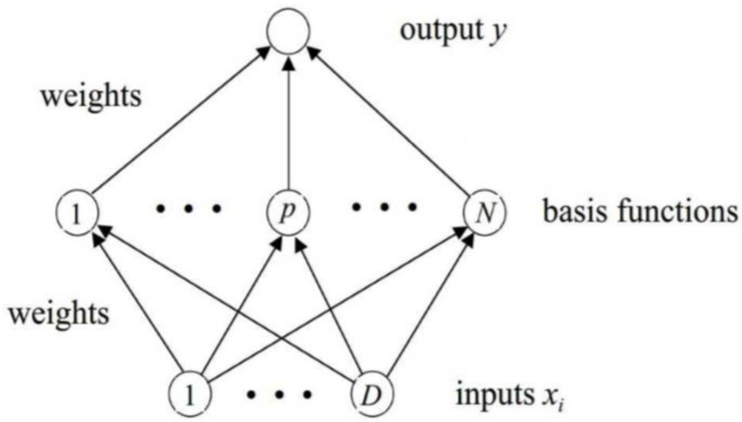


**Figure 4. RBF Network Structure**

## 4. Support Vector Machines

Support Vector Machines has been introduced by Vapnik and his colleagues [26], SVM models are very similar to classical multilayer perceptron neural networks used for classification [27], but recently they have been extended to solve regression problems [28]. Moreover, SVM has been shown a success use in some areas such as; Intrusion Detection [29], Handwritten Recognition [30], Face Detection [31], Time Series Prediction [32] and others.

SVM is very similar to an ANN since both receive input data and provide output data. For regression, the input and output of SVM are identical to the ANN. However, what makes the SVM primarily better is that the SVM does not suffer from over fitting like ANN does. So, the ANN memorizes the input data on the training stage and will not perform well at the testing data. The kernel function that has been used for the SVM in this work is Radial Basis Function (RBF) as the following equation:

$$K\left(x, y\right) = \exp\left(-\|x - y\|^2 / 2\delta^2\right) \qquad (4)$$

## 5. Model Evaluation

In order to check the performance of the developed ANN models, the Mean Squares Error (MSE), the Euclidian distance (ED), the Manhattan distance (MD) and the Mean Magnitude of Relative Error (MMRE) were measured. These performance criteria are assessed to measure how close the measured values to the values developed using the ANN approach. MSE, ED and MD are computed as:

1) Mean Squares Error (MSE)

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y - \hat{y})^2 \qquad (5)$$

2) Euclidian distance (ED):

$$ED = \sqrt{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \qquad (6)$$

**Table II.**

Parameters setting for FF-ANN

| FF-ANN Parameter | Value |
|---|---|
| Number of neurons in hidden layer | 2 |
| Learning Rate | 0.3 |
| Momentum | 0.2 |
| Number of Epochs | 2000 |

3) Manhattan distance (MD):

$$MD = (\sum_{i=1}^{n} |y_i - \hat{y}_i|) \qquad (7)$$

4) Mean Magnitude of Relative Error (MMRE)

$$MMRE = \frac{1}{n}\sum_{i=1}^{n} \frac{|y_i - \hat{y}_i|}{y_i} \qquad (8)$$

where y and $\hat{y}$ are the actual lipase activates and the estimated activates based on proposed model and *n* is the number of measurements used in the experiments, respectively.

## 6. Experimental Results

The definitive objective of our experiments in this work is to find the values of the Surface Ozone layer produced by each model of a function of Nitrogen-di-oxide, temperature and Relative Humidity. Therefore, our goal is to build three models structures which each one has multiple inputs and single output as given in Table I. WEKA framework was used to apply the three models in the experiments designed in this research. It is open source software implemented in JAVA and developed at the University of Waikato, New Zealand. WEKA provides a set of machine learning algorithms for data mining tasks. (http://www.cs.waikato.ac.nz/ml/weka/). For the ANNs guidelines, Table II shows the parameters for the FF-ANN chosen after some primary experiments and Table III shows the parameters for the RBF-ANN. Moreover, the data set was divided into two parts. The first 75% of the data set was used in the training phase of the ANN modeling process while the 25% of the second data set was used for the testing phase.

MSE, ED, MD and MMRE performance criterion were used to measure the prediction results against the observed values as they verified in the previous section to see the most stable and suitable model for the Ozone Layer. Finally, the predicted values are compared with results obtained by each model. The results in the training and testing stage are given in Table IV and Table V respectively. It can be seen clearly that the performance of the prediction capability of RBF-ANN has a superior prediction power compared to the FF-ANN and SVM. Moreover, the SVM is also much better than the FF-ANN model. Figure 5, Figure 6 and Figure 7 illustrates the performance of the FF, RBF and SVM model over the testing data set for the Ozone level respectively.

**Table II.**

Parameters setting for FF-ANN

| RBF-ANN Parameter | Value |
|---|---|
| Minimum Standard Deviation | 0.1 |
| Ridge value for the logistic | 1.0e-8 |

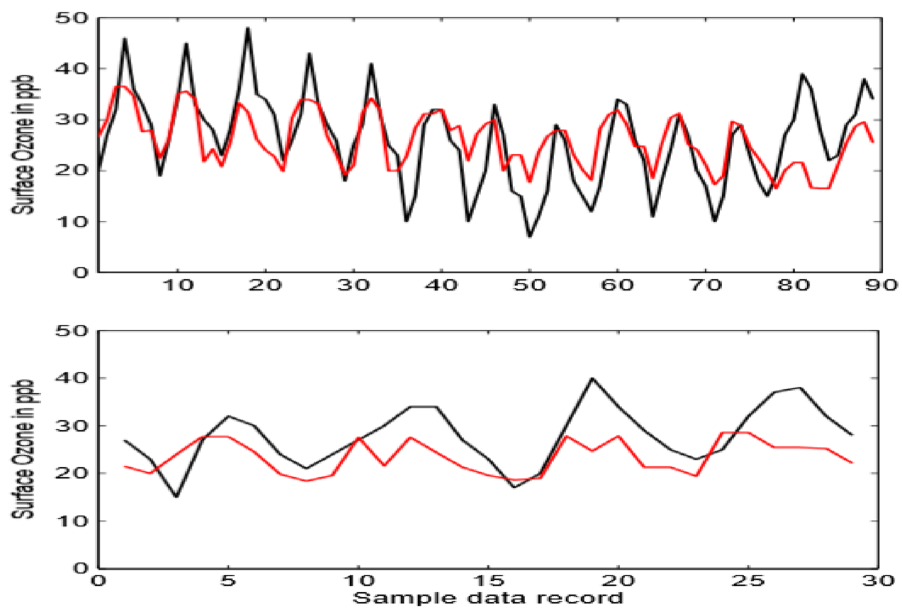


**Figure 5. Actual and Estimated Surface Ozone Measurements using FF-ANN Model in Training and Testing Cases**

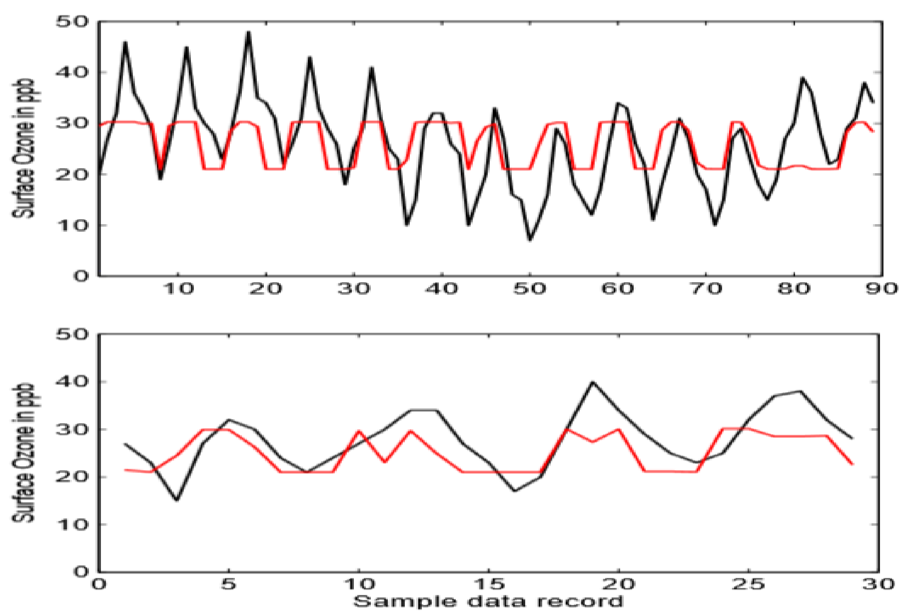


**Figure 6. Actual and Estimated Surface Ozone Measurements using RBF- ANN Model in Training and Testing Cases**

**Table II.**

Parameters setting for FF-ANN

| RBF-ANN Parameter | Value |
|---|---|
| Minimum Standard Deviation | 0.1 |
| Ridge value for the logistic | 1.0e-8 |

**Table IV.**

Computer MSE, ED, MD AND MMRE in the training cases

| | MSE | ED | MD | MMRE |
|---|---|---|---|---|
| FF-ANN | 46.734 | 64.493 | 5.415 | 0.26599 |
| RBF-ANN | 56.788 | 71.092 | 5.9805 | 0.29528 |
| SVM | 47.979 | 65.346 | 5.3803 | 0.27339 |

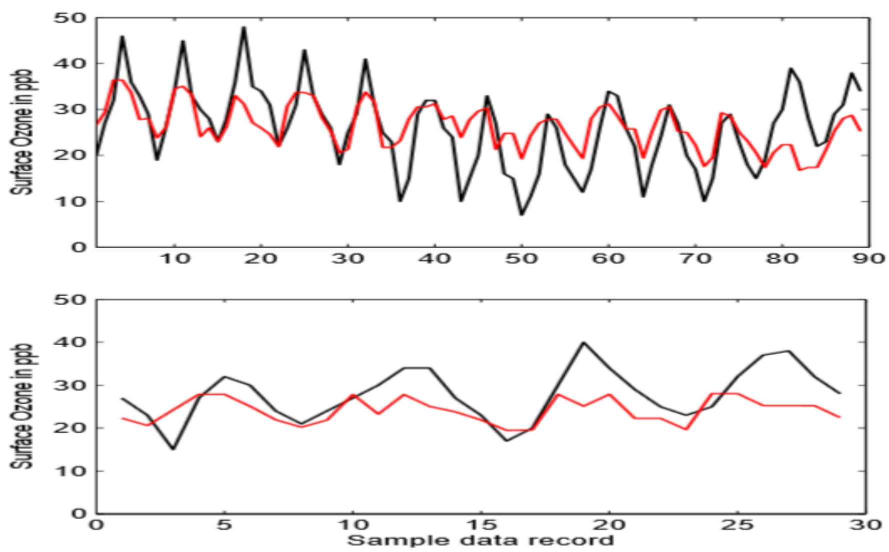


**Figure 7. Actual and Estimated Surface Ozone Measurements using Support Vector Machines (SVM) Model in the Testing Case**

**Table V.**

Computed MSE, ED, MD AND MMRE in the testing cases

| | MSE | ED | MD | MMRE |
|---|---|---|---|---|
| FF-ANN | 51.197 | 67.502 | 5.8173 | 0.29503 |
| RBF-ANN | 47.275 | 64.865 | 5.4974 | 0.289 |
| SVM | 49.544 | 66.404 | 5.5974 | 0.29016 |

## 7. Conclusion and Future Work

The main objective of the presented study is to make comparability between distinguished neural network models and SVM to see what is the best type for Ozone level prediction to have an early and accurate alert. Experiment and analysis were conducting to point out some advantages of RBF neural network compared to other FF ANN model and SVM. The developed RBF model provided good estimation and prediction capabilities in training and testing cases as it has been shown on the result's section. In the future, the research should focus on exploring other advantages of RBF and make further analysis.

## References

[1] C. Varotsos, C. Tzanis, S. Tsitomeneas, M. N. Assimakopoulos and A. Mammis, "Surface solar ultraviolet irradiance and total ozone during summertime", Int. J. Remote Sens., vol. 29, **(2008)** May, pp. 2667-2673.

[2] M. Norval, R. Lucas, A. Cullen, F. De Gruijl, J. Longstreth, Y. Takizawa and J. Van Der Leun, "The human health effects of ozone depletion and interactions with climate change", Photochemical & Photobiological Sciences, vol. 10, no. 2, (**2011)**, pp. 199-225.

[3] A. Webb, "Health hazards of ozone depletion", Weather, vol. 44, no. 5, **(2012)**, pp. 215-220.

[4] G. Dixon, "Climate change–impact on crop growth and food production, and plant pathogens", Canadian Journal of Plant Pathology, vol. 34, no. 3, **(2012)**, pp. 362-379.

[5] Y. Uchiyama, R. Fujisawa, Y. Oda and E. Hirasawa, "Air conditioner and washing machine primary disassembly process", Proceedings of the First international conference on Environmentally conscious design and inverse manufacturing, ecodesign'99, (Washington, DC, USA), IEEE Computer Society, **(1999)**, pp. 258-262.

[6] H. Jing, "Preparation of tio2/acf in caprolactam tetrabutyl ammonium bromide ionic liquid", Proceedings of the 2012 International Conference on Biomedical Engineering and Biotechnology, ICBEB '12, (Washington, DC, USA), IEEE Computer Society, **(2012)**, pp. 1610-1613.

[7] A. Woodcock, "Prevention is best: lessons from protecting the ozone layer", Thorax, **(2012)**.

[8] A. Sheta, "Software effort estimation and stock market prediction using takagisugeno fuzzy models", Proceedings of the 2006 IEEE Fuzzy Logic Conference, Sheraton, Vancouver Wall Centre, Vancouver, BC, Canada, **(2006)** July 16-21, pp. 579-586.

[9] A. Sheta and K. Eghneem, "Training artificial neural networks using genetic algorithms to predict the price of the general index for amman stock exchange", Proceedings of the Midwest Artificial Intelligence and Cognitive Science Conference, DePaul University, Chicago, IL, USA, **(2007)** April 21-22, pp. 7-13.

[10] M. Moustra, M. Avraamides and C. Christodoulou, "Artificial neural networks for earthquake prediction using time series magnitude data or seismic electric signals", Expert Syst. Appl., vol. 38, November **(2011)**, pp. 15032-15039.

[11] J. Thibault, V. V. Breusegem and A. Cheruy, "Online prediction of fermentation variables using neural networks", vol. 36, **(1990)**, pp. 1041-1048.

[12] S. S. Fan, "Problems of stream flood simulation and prediction", Proc. of Hydraulic Engineering, **(1994)**, pp. 1110-1114.

[13] A. Sheta and A. Mahmoud, "Forecasting using genetic programming", Proceedings of the 33rd Southeastern Symposium on System Theory, **(2001)**, pp. 343-347.

[14] Y. Al-Zubi, A. Sheta and J. Al-Zubi, "Nile river flow forecasting based takagisugeno fuzzy model", Journal of Applied Sciences, vol. 10, **(2010)**, pp. 284-290.

[15] U. Kumar and V. Jain, "Arima forecasting of ambient air pollutants (o 3, no, no 2 and co)", Stochastic Environmental Research and Risk Assessment, vol. 24, no. 5, **(2010)**, pp. 751-760.

[16] M. Boznar, M. Lesjak and P. Mlakar, "A neural network-based method for short-term predictions of ambient so¡ sub¿ 2¡/sub¿ concentrations in highly polluted industrial areas of complex terrain", Atmospheric Environment. Part B. Urban Atmosphere, vol. 27, no. 2, **(1993)**, pp. 221-230.

[17] V. Uyak, O. Ozkan, I. Koyuncu and K. Demirbas, "Modeling air pollutants in turkey with artificial neural network and multiple lineer regression", **(2012)**.

[18] R. Selvaraj, R. Gayathri, K. Elampari and S. Jeyakumar, "A neural network model for short term prediction of surface ozone at tropical city", International Journal of Engineering Science, vol. 2, **(2010)**.

[19] W. McCulloch and W. Pitts, "A logical calculus of the ideas immanent in nervous activity", Bulletin of mathematical biology, vol. 5, no. 4, **(1943)**, pp. 115-133.

[20] H. Hruschka, "Determining market response functions by neural networks modeling: a comparison to econometric techniques", European Journal of Operation Research, vol. 66, **(1993)**, pp. 867-888.

[21] K. Chakraborty, "Forecasting the behaviour of multivarable time series using neural networks," Neural Networks, vol. 5, **(1992)**, pp. 962-970.

[22] E. Y. Li, "Artificial neural networks and their business applications", Information and Managements, vol. 27, **(1994)**, pp. 303-313.

[23] G. Swales and Y. Yoon, "Applying artificial neural networks to investment analysis", Finacial Analyst Jornal, **(1992)** September-October.

[24] B. Krse and P. van der Smagt, "An Introduction to Neural Networks", CRC Press, The University of Amsterdam, **(1993)**.

[25] L. V. Fausett, "Fundamentals of Neural Networks: Architectures, Algorithms, and Applications", Prentice Hall, **(1994)**.
[26] C. Cortes and V. Vapnik, "Supportvector networks", Machine learning, vol. 20, no. 3, **(1995)**, pp. 273-297.
[27] R. Hua, "Dai liankui (state key lab of industrial control technology, zhejiang university, hangzhou 310027, china); support vector machine classification and regression based hybrid modeling method and its application in raman spectral analysis", Chinese Journal of Scientific Instrument, vol. 11, **(2010)**.
[28] V. Vapnik, S. Golowich and A. Smola, "Support vector method for function approximation, regression estimation, and signal processing", Advances in neural information processing systems, **(1997)**, pp. 281-287.
[29] D. Hong, "Application of support vector machine in intrusion detection", Computer Engineering, vol. 4, **(2012)**, pp. 047.
[30] O. Rashnodi, H. Sajedi, M. Abadeh, A. Elci, M. Munot, M. Joshi, N. Sharma, N. Gupta, R. Sharma and B. Mihajlov, "Persian handwritten digit recognition using support vector machines", International Journal of Computer Applications, vol. 29, no. 12, **(2011)**, pp. 1-6.
[31] J. Ruan and J. Yin, "Face detection based on facial features and linear support vector machines", Communication Software and Networks, 2009. ICCSN'09. International Conference on, IEEE, **(2009)**, pp. 371-375.
[32] N. Sapankevych and R. Sankar, "Time series prediction using support vector machines: a survey", Computational Intelligence Magazine, IEEE, vol. 4, no. 2, **(2009)**, pp. 24-38.

## Authors

**Dr. Mouhammd Alkasassbeh**, BSc, MSc, MPhil, PhD, is assistance professor at Mutah University specialized in computer networks. His research interest includes computer networking, sensor network applications, artificial neural network, genetic programing and also bioinformatics.